AD_____

Award Number: W81XWH-04-1-0197

TITLE: Genomic Approaches for Detection and Treatment of Breast Cancer

PRINCIPAL INVESTIGATOR: Stephen J. Elledge, Ph.D.

CONTRACTING ORGANIZATION: Brigham Womens Hospital Inc
Boston MA 02115-6110

REPORT DATE: July 2006

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| **1. REPORT DATE** *(DD-MM-YYYY)* | **2. REPORT TYPE** | **3. DATES COVERED** *(From - To)* |
|---|---|---|
| 01-07-2006 | Annual | 01 Jul 05 – 30 Jun 06 |

**4. TITLE AND SUBTITLE**
Genomic Approaches for Detection and Treatment of Breast Cancer

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
W81XWH-04-1-0197

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
Stephen J. Elledge, Ph.D.

E-Mail: selledge@genetics.med.harvard.edu

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

Brigham Womens Hospital Inc
Boston MA 02115-6110

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**
None provided.

**15. SUBJECT TERMS**
None provided.

| **16. SECURITY CLASSIFICATION OF:** | | | **17. LIMITATION OF ABSTRACT** | **18. NUMBER OF PAGES** | **19a. NAME OF RESPONSIBLE PERSON** |
|---|---|---|---|---|---|
| **a. REPORT** | **b. ABSTRACT** | **c. THIS PAGE** | | | USAMRMC |
| U | U | U | UU | 23 | **19b. TELEPHONE NUMBER** *(include area code)* |

# Table of Contents

**Introduction**

A key part of our research plan has been the development and use of retroviral vectors expressing RNA interference RNAs to identify human genes involved in causing or restraining cancer. In our first progress report we described our efforts to develop shRNA libraries and showed they could be used to identify tumor suppressors. Ultimately our goal is to screen of complex pools of shRNA expressing retroviruses each marked with a bar code that allows the results of the screen to be read out by microarray hybridization. We demonstrated this could be accomplished in enrichment screens for shRNAs that caused cellular transformation and growth in soft agar. However, a key goal has been to identify shRNAs that debilitate or kill cancer cells. In order for this to be possible in complex pools, it is imperative that each vector knock down its target with high penetrance. We have successfully achieved this level of knockdown and can now see particular shRNA expressing viruses drop out of complex pools. We are also developing bar code hybridization methods that allow us to detect over 80% of the viruses in pool. We hope to push this to over 90% using double bar codes. Together the high knockdown vectors together with the bar code hybridization has allowed us to achieve the central goal of this Award.

We have taken advantage of our improvements to perform a variety of screens relevant to cancer. These include lethality screens to identify genes that cancer cell lines rely upon to proliferate and survive. We also are performing screens that are identifying genes whose loss gives rise to resistance to tyrosine kinase inhibitors. Finally we are continuing our efforts to identify tumor suppressors in human mammary epithelial cells. We have expanded this to identify oncogenes in the same system.

In parallel we are attempting to generate a system through which we can explore the autoimmune phenotype of breast cancer patients. We have generated the peptide display libraries required for this and hope to have some screening information during the next installment of this Award. Below are more detailed descriptions of our efforts.

## Body

**Progress on barcode screening for essential genes:**

To test if barcode screens can work robustly to quickly and efficiently identify genes that are required for cell survival and proliferation, we performed a pilot screen in a well-characterized human cancer cell line Hct116. In the screen, we infected Hct116 cells with one of our retroviral microRNA-based shRNA libraries containing 13690 shRNAs. We split the infected cells into 2 halves 48 hours after infection and collected half of the cells as a reference sample. The other half of the cells was continuously propagated in culture for additional three weeks and cells were collected from two different time points during the culture (day14 and day21) as experimental samples. We predict that shRNAs causing cell death or cell cycle arrest will gradually disappear during the prolonged period of culture, while other shRNAs will remain largely unchanged. To measure the change in abundance for each shRNA in the cells over time, we isolated genomic DNA from both the reference and experiment samples and PCRed the shRNAs out of the genomic DNA. We labeled the shRNAs PCRed from the reference sample with one color

(eg., Cy5-labled nucleotides) and the experiment samples with another color (eg., Cy3-labled nucleotides), and hybridized them on microarrays that contain probes against every individual shRNA in the library. As shown in the following figure A-C, we found many shRNAs that had reduced abundance over time, suggesting that the genes they targeting may be required for normal cell proliferation. In 3 independent trials, we identified a total of 218 shRNAs that were consistently dropping out during

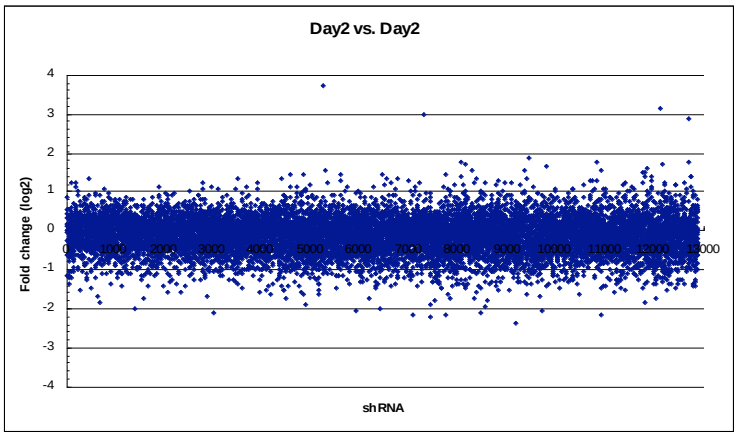
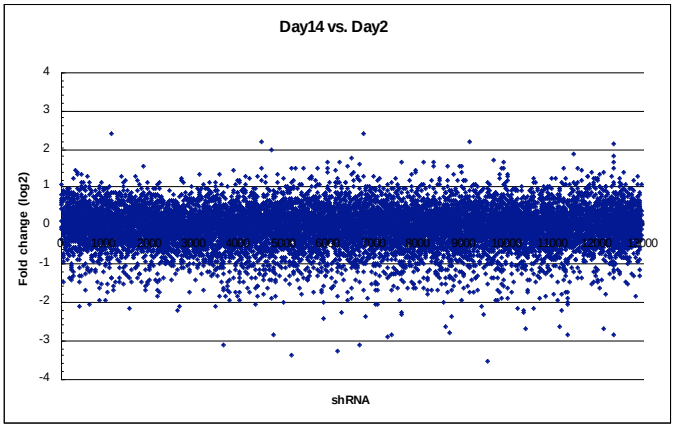



**Fig. 1 A.**
Hybridization of barcodes from the independent samples 2 days after viral infection. Very little variance is observed.

the culture (Figure 1D). We are currently cherry-picking the candidate shRNAs and we will study the effect of these shRNAs and the genes they target individually on cell survival and proliferation.

Our current data suggested that the barcode screen can be a very efficient and cost-friendly method to screen for genes that are essential for cell survival and proliferation. We are currently applying the same screening methodology to the breast cancer cell lines and breast cancer primary cells to look for genes that are required for the growth of these cancer.

**Fig. 1B**
Hybridizat of barcodes from the samples 2 days after viral infection compared to samples 14 days after infection. Note there is more variance in these samples than in 1A as would be expected if genes are selected against.



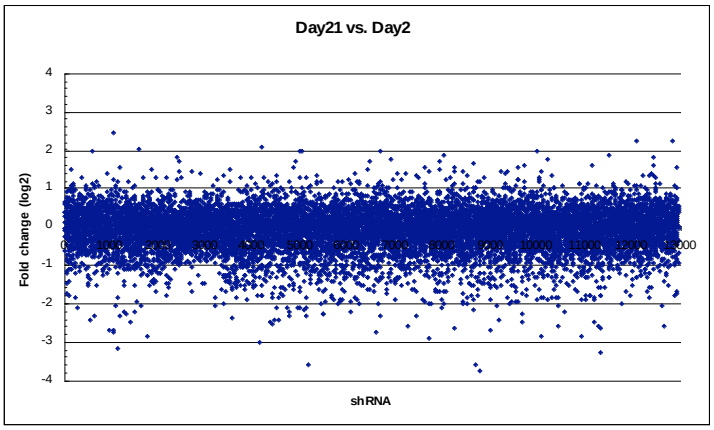



**Fig. 1C**
Hybridization of barcodes from the samples 2 days after viral infection compared to samples 21 days after infection. Note there is even more variance in these samples than in 1B as would be expected if genes are selected against.

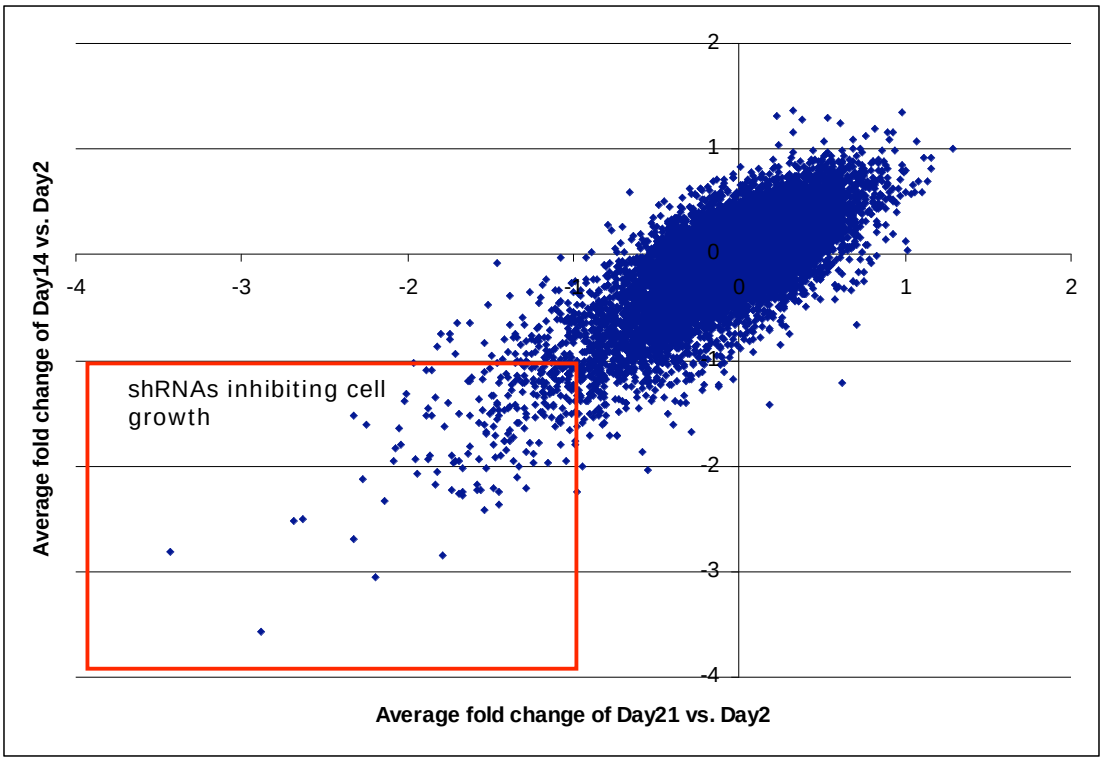



**Fig. 1D**
When you plot the average fold change of Day 2 vs 14 against Day2 vs Day 21, the shRNAs that continue to drop out of the population in both experiments show up in the lower left quandrant of the graph. These, shown in a red box, are the candidates for lethal genes in tumor cells and represent potential anti-cancer drug targets.

**Identification of genes that confer acquired resistance to EGFR inhibitors**

The ErbB family of receptor tyrosine kinases represents an important class of drug targets in cancer. A significant fraction of breast tumors overexpress ErbB2 (HER2) and ErbB3, whereas amplification and activating mutations in ErbB1, also known as the epithelial growth factor receptor (EGFR), have been found in lung cancers (Hynes and Lane, 2005). Targeted therapies against ErbB receptors have shown great promises in treating tumors addicted to ErbB signaling. This is exemplified by the application of the inhibitory antibody trastuzumab (Herceptin) in the treatment of $HER2^{+}$ breast cancers and by the application of small-molecule EGFR inhibitors gefitinib (Iressa) and erlotinib (Tarceva) in the treatment of non-small cell lung cancers (NSCLCs) harboring mutant EGFR.

Despite their initial efficacies, tumors rapidly develop acquired resistance upon treatment, thus severely hampering the long-term benefit of these agents. A better understanding of the mechanisms underlying to such acquired resistance, therefore, would provide new opportunities for combination and second-line therapeutics.

We chose the NSCLC cell line HCC827 as a model system to study acquired resistance to EGFR inhibitors. As EGFR and ErbB2 share similar downstream signaling pathways (such as the MAP kinase and PI 3-kinase pathways), the knowledge gained from this study should be generally applicable to tumors of the breast and other organs that are driven by ErbB receptor signaling.

HCC827 cells harbor activating mutations in EGFR and are exquisitely sensitive to the EGFR inhibitor gefitinib (Mukohara et al., 2005) with an IC50 (50% inhibitory concentration) at ~5nM (Figure 1A). We are currently conducting a genome-wide loss-of-function screen using our mir30-based shRNA library (Silva et al., 2005) to identify genes whose loss of function would confer resistance to gefitinib in these cells. Briefly, pools of retroviral shRNAs (~12,000 distinct shRNAs per pool) were stably introduced into cells at high representation (>500 copies per shRNA) and low multiplicity of infection (0.2 virus per cell) to create stable pools of cells in which most cells contains only a single shRNA. Each pool of cells was divided into two halves, with one half propagated for 10 doublings in regular growth media (control group) and the other 10 doublings in growth media containing 5nM gefitinib (gefitinib group, Figure 1B). We chose this particular concentration of gefitinib as it inhibits cell growth by ~50% (Figure 1C), therefore allowing us to identify both shRNAs that confer gefitinib resistance and shRNAs that enhances gefitinib sensitivity. For each group shRNAs were PCR recovered from the initial and end samples and the change in relative representation of a particular shRNA over time in each group was measured by competitive hybridization to a microarray (Figure 1B). The experiment was carried out in triplicates and shRNAs that are consistently enriched in the gefitinib group but not in the control group are candidates that confer gefitinib resistance. Conversely, those shRNAs that consistently dropped out from the gefitinib group but not from the control group are candidates that enhance gefitinib sensitivity.

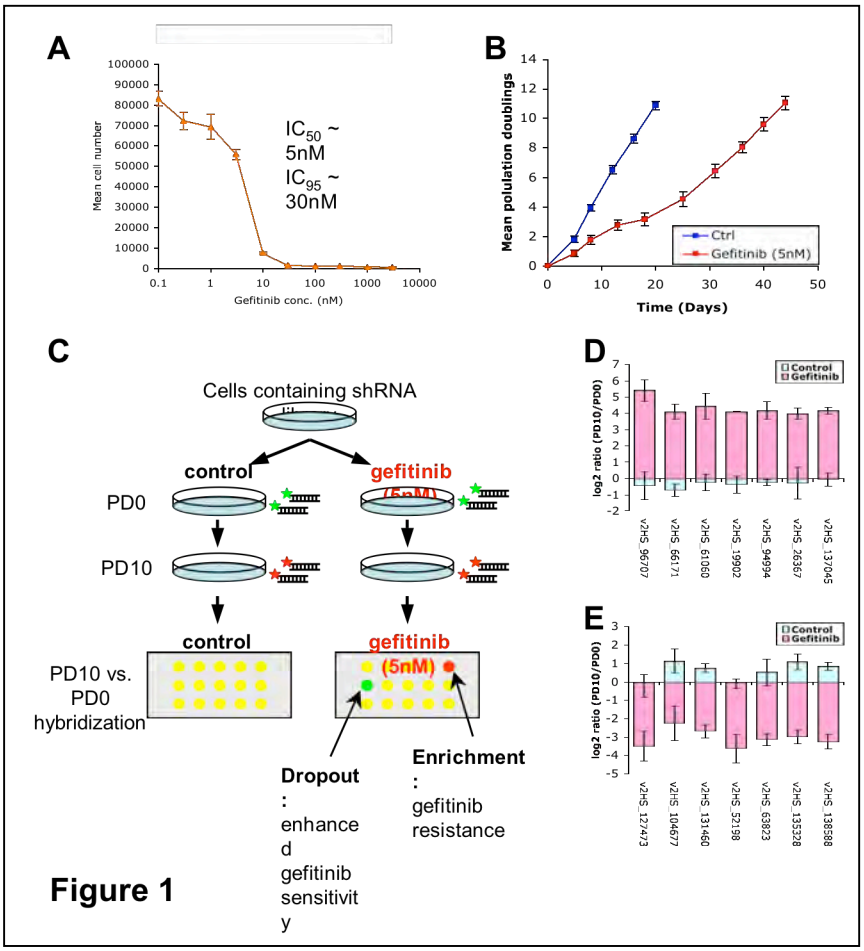



Figure 1

**Figures 1D and 1E** shows examples of shRNAs from one such screen. We are currently validating individual shRNAs from the screen. The fact that we observe multiple shRNAs against the same gene scoring in the screen lent further support to the validity of our technology platform of screening shRNAs in large pools followed by microarray deconvolution.

**Figure 1.** Gefitinib resistance screen. **A.** Gefitinib sensitivity curve of HCC827 cells. **B.** Cell growth under control condition and in the presence of 5nM gefitinib. This particular concentration of gefitinib was chosen such that both shRNAs that confer resistance and those that confer enhanced sensitivity can be identified in the screen. **C.** A schematic layout of the screen. Stable pools of cells containing the same shRNA library were split into two halves and grown either in normal media (control) or in the presence of 5nM gefitinib for 10 population doublings (PD). Genomic DNA samples were collected at population doubling 0 (PD0) and 10 (PD10). After PCR-recovery of shRNAs, for each condition the PD10 and PD0 samples were competitively hybridized to a bar-code microarray. ShRNAs that confers gefitinib resistance will be enriched over time in the gefitinib group but not in the control group, whereas those that confer enhanced gefitinib sensitivity will drop out over time in the gefitinib group but not in the control group. **D.** Examples of shRNAs from the screen that confer gefitinib resistance. **E.** Examples of shRNAs that confer enhanced gefitinib sensitivity. In **D** and **E** the log2 ratio of the shRNAs relative abundance between the PD10 and PD0 samples are plotted for each condition (mean ± s.d.).

## Identification of genes whose knockdown confers synthetic lethality with the Ras oncogene

An attractive approach in cancer therapeutics is to exploit the genetic differences between cancer cells and normal cells. Tumors are often driven by a handful of oncogenic mutations such as Ras mutation, myc overexpression, p53 loss or PTEN loss. Identifying genes whose loss of function constitute synthetic lethality with such mutations would provide new drug targets that allow the selective killing of cancer cells while sparing normal cells.

We chose oncogenic Ras as a model system as Ras is one of the most frequently mutated oncogenes. To identify genes whose loss of function constitute synthetic lethality with oncogenic Ras, we used the isogenic colorectal cancer cell lines DLD-$1^{Mut}$ and DLD-$1^{WT}$ for a mir30 shRNA-based loss-of-function screen. We would have used a breast line but none were isogenic for ras. The parental DLD-$1^{Mut}$ cells harbor an activating mutation in one of the two K-ras alleles. This mutant K-ras allele was somatically knocked out in the derivative DLD-$1^{WT}$ cells. As a result, DLD-$1^{WT}$ cells have lost the capacity for anchorage-independent growth and show reduced proliferation rate on solid substrate (Shirasawa et al., 1993). We introduced the same pool of retroviral shRNAs (~12,000 distinct shRNAs per pool) into both DLD-1Mut and DLD-1WT

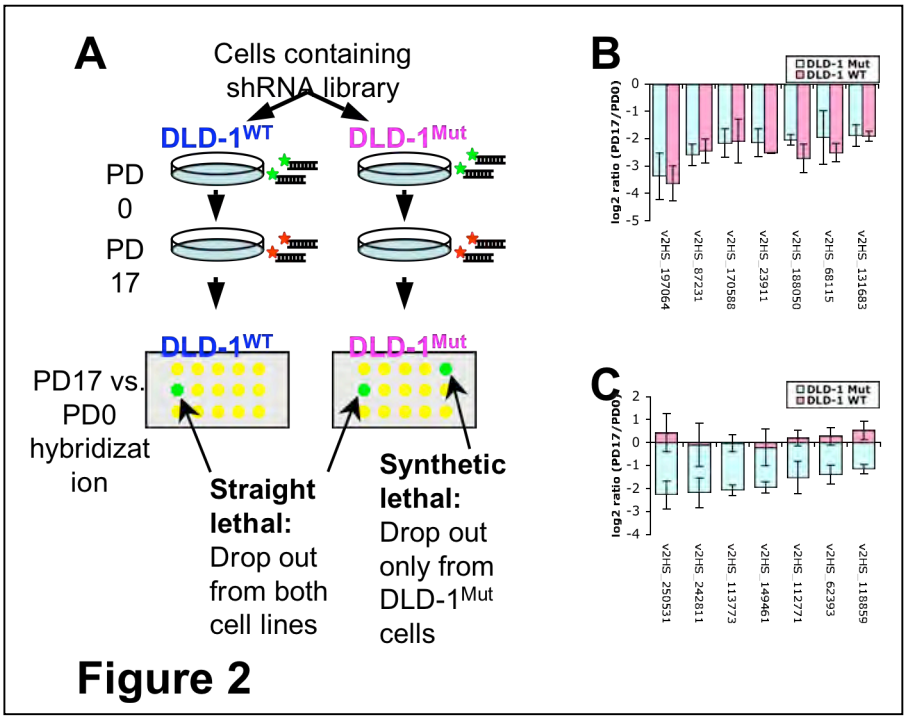



**Figure 2**

**Figure 2.** Ras synthetic lethal screen. **A.** Schematic. The isogenic pair of cell lines, DLD-$1^{WT}$ and DLD-$1^{Mut}$, which differ only by the presence of a mutant K-ras were infected with the shRNA library, propagated for 17 population doublings (PD). After PCR-recovery of bar codes the PD17 and PD0 samples were competitively hybridized to a bar-code microarray. shRNAs that are lethal to both cell lines will drop out in both cases over time, whereas those that are synthetically lethal with mutant K-ras will only drop out from the DLD-$1^{Mut}$ cells but not from the DLD-$1^{WT}$ cells. **B.** Straight lethal shRNAs that dropped out from both lines. **C.** shRNAs that dropped out only from DLD-$1^{Mut}$ cells and are candidates that confer synthetic lethality with mutant K-ras. In **B** and **C** the log2 ratio of the shRNAs relative abundance between the PD10 and PD0 samples are plotted for each condition (mean ± s.d.).

cells at high representation and low multiplicity and propagated the resulting cells for 17 doublings. The shRNAs within each cell line were PCR recovered and changes in their relative abundances at the end vs. the beginning of the experiment were measured by microarray hybridization (Figure 2A). Those shRNAs that target essential genes and thus impair the growth of both cell lines would drop out from both lines over time (Figure 2B). On the other hand, shRNAs that specifically impair the viability of DLD-$1^{Mut}$ but not DLD-$1^{WT}$ cells will only dropout from the former and not from the latter (Figure 2C). It is this class of shRNAs that are synthetically lethal with mutant K-ras we are particularly interested in as their target genes represent attractive therapeutic targets for treating tumors bearing activating K-ras mutations. We are currently in the process of validating these shRNA hits and expanding the screen to genome-scale. Once we identify promising hits and validate them, we will survey other cancer lines, especially breast cancer lines with mutant ras to determine the generality of our findings.

## References


Hynes, N. E., and Lane, H. A. (2005). ERBB receptors and cancer: the complexity of targeted inhibitors. Nat Rev Cancer *5*, 341-354.

Mukohara, T., Engelman, J. A., Hanna, N. H., Yeap, B. Y., Kobayashi, S., Lindeman, N., Halmos, B., Pearlberg, J., Tsuchihashi, Z., Cantley, L. C., Tenen, D. G., Johnson, B. E., and Janne, P. A. (2005). Differential effects of gefitinib and cetuximab on non-small-cell lung cancers bearing epidermal growth factor receptor mutations. J Natl Cancer Inst *97*, 1185-1194.

Shirasawa, S., Furuse, M., Yokoyama, N., and Sasazuki, T. (1993). Altered growth of human colon cancer cell lines disrupted at activated Ki-ras. Science *260*, 85-88.

Silva, J. M., Li, M. Z., Chang, K., Ge, W., Golding, M. C., Rickles, R. J., Siolas, D., Hu, G., Paddison, P. J., Schlabach, M. R., Sheth, N., Bradshaw, J., Burchard, J., Kulkarni, A., Cavet, G., Sachidanandam, R., McCombie, W. R., Cleary, M. A., Elledge, S. J., and Hannon, G. J. (2005). Second-generation shRNA libraries covering the mouse and human genomes. Nat Genet *37*, 1281-1288.

## Screens for Tumor suppressors using the RNAi library

In our previous progress report, we described our effort to identify candidate human tumor suppressors by combining an *in vitro* model of human mammary epithelial cell (HMEC) transformation and our first-generation RNAi library (Westbrook et al., 2005). Recently, we have investigated the molecular mechanisms governing the function of REST, one of the novel tumor suppressors identified in our initial screen. We have also expanded the screen by employing a more potent fourth-generation shRNA library targeting all human kinases and phosphatases. In addition, we have utilized an ORFeome-based cDNA library in the context of the same system to identify potential oncogenes.

## Identification and characterization of a novel REST ubiquitin ligase

The success of the initial screen is emphasized by our discovery of known human tumor suppressors as well as genes with unappreciated roles in human cancer. In this regard, we have continued to explore the physiologic and pathologic roles of one of these novel candidates, REST. REST, or RE1-silencing transcription factor, was originally discovered as a transcriptional repressor of neuronal genes in non-neuronal tissues (Chong et al., 1995; Schoenherr and Anderson, 1995). Our previous experiments uncovered REST as a potent suppressor of cellular transformation and PI3K signaling, a pathway well-documented for its role in human cancers (Westbrook et al., 2005). Our more recent studies have focused on the upstream regulation of REST function. More specifically, we have shown that REST is regulated at the level of protein turnover. We have shown that REST is ubiquitinated *in vivo* and, using a genetic screen, we have discovered the ubiquitin ligase that facilitates REST degradation. Consistent with its role as a regulator of REST function, this ubiquitin ligase governs neural differentiation and mammary epithelial cell transformation (Figure 1). This ubiquitin ligase requires REST-phosphorylation, indicating that a signal transduction network controls REST degradation. Such a network regulating REST function would have important physiologic and pathologic significance, and we are currently using genetic methods to elucidate the components of this network.

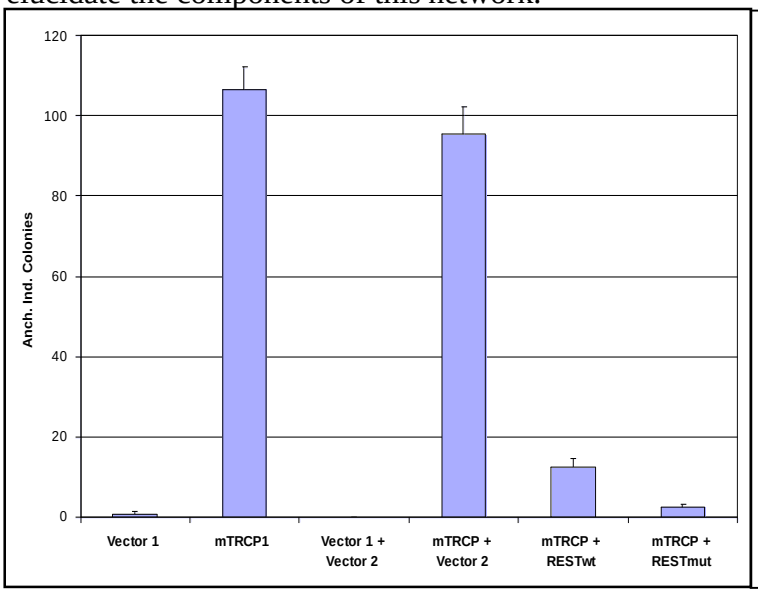



Figure-1: TRCP-mediated transformation of HMECs requires loss of REST function. HMECs were transduced with retroviruses expressing the indicated cDNAs and assessed for anchorage independent proliferation. RESTmut indicates a REST cDNA with 3 point mutations in the TRCP-binding domain.

Expanded screens for candidate tumor suppressors and oncogenes

For our initial screen, we employed our first-generation shRNA library encompassing 20,000 shRNAs targeting 9,000 human genes. However, there have been several advances in our understanding of shRNA processing and miRNA biogenesis since the construction of this library, and we have incorporated these insights into subsequent generations of our shRNA libraries (Silva et al., 2005). Consequently, these libraries elicit more efficient knockdown of gene targets and should allow a more complete interrogation of the human genome. Consequently, we expanded the initial HMEC-transformation screen by using a fourth-generation shRNA library targeting all human kinases and phosphatases (roughly 4200 shRNAs targeting 700 genes). Notably, this screen identified roughly the same number of candidates (22 genes) as the initial screen, though this library targeted 13-fold fewer genes. This is consistent with the improved penetrance of the currently library-design. Notably, we isolated PI3K-mTOR pathway components that have previously been implicated in tumor suppression (eg. PTEN, LKB1). Intriguingly, recent genome sequencing efforts have described somatic mutations in many of these candidates in various malignancies. Our identification of such genes in a functional genetic screen suggests that their dysregulation may contribute to human cancer. Moreover, our cell-based systems provide us a context in which to study such tumor-derived mutants and their phenotypes (see Westbrook et al., 2005). In preliminary experiments, we have begun to delineate the specific consequences of candidate dysfunction on mammary morphogenesis. Using 3-D *in vitro* cultures of mammary epithelial cells (Debnath and Brugge, 2005; Debnath et al., 2002; Muthuswamy et al., 2001), we can simultaneously probe the effects of candidate loss-of-function on proliferation, survival, and cell polarity. For example, shRNAs targeting a candidate phosphatase elicit aberrant mammary acini (Figrue 2). This approach will enable us to begin to construct tumor suppressor networks based on the candidates' cell biological functions.

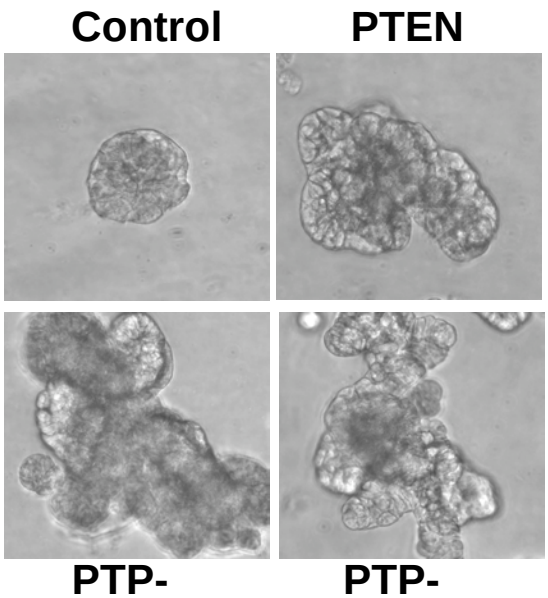



Figure 2. Candidate PTP function is required for proper mammary acinar morphogenesis. MCF10A mammary epithelial cells were transduced with lentiviruses expressing shRNAs that target the indicated proteins. Cells were cultured in matrigel for 15 days, and observed under phase contrast microscopy. Note: 2 shRNAs targeting the candidate PTP yielded similar results.

In addition to these loss-of-function screens, we have interrogated the HMEC transformation system with an ORFeome-based cDNA library. This library, which was constructed by subcloning the ORFeome version-1 collection into a retroviral vector, is advantageous over standard cDNA libraries because of its size (> 8,000 cDNAs) and normalization. Using this tool, we discovered > 80 candidate oncogenes, consisting of both known oncogenes (eg. NTRK3) as well as genes regulating proto-oncogenic pathways (eg. WNT signaling). We are currently validating these candidates, focusing on those genes involved in pathways or processes not previously implicated in human cancer.

**References**


Chong, J. A., Tapia-Ramirez, J., Kim, S., Toledo-Aral, J. J., Zheng, Y., Boutros, M. C., Altshuller, Y. M., Frohman, M. A., Kraner, S. D., and Mandel, G. (1995). REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. Cell *80*, 949-957.

Debnath, J., and Brugge, J. S. (2005). Modelling glandular epithelial cancers in three-dimensional cultures. Nat Rev Cancer *5*, 675-688.

Debnath, J., Mills, K. R., Collins, N. L., Reginato, M. J., Muthuswamy, S. K., and Brugge, J. S. (2002). The role of apoptosis in creating and maintaining luminal space within normal and oncogene-expressing mammary acini. Cell *111*, 29-40.

Muthuswamy, S. K., Li, D., Lelievre, S., Bissell, M. J., and Brugge, J. S. (2001). ErbB2, but not ErbB1, reinitiates proliferation and induces luminal repopulation in epithelial acini. Nat Cell Biol *3*, 785-792.

Schoenherr, C. J., and Anderson, D. J. (1995). The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. Science *267*, 1360-1363.

Silva, J. M., Li, M. Z., Chang, K., Ge, W., Golding, M. C., Rickles, R. J., Siolas, D., Hu, G., Paddison, P. J., Schlabach, M. R., *et al.* (2005). Second-generation shRNA libraries covering the mouse and human genomes. Nat Genet *37*, 1281-1288.

Westbrook, T. F., Martin, E. S., Schlabach, M. R., Leng, Y., Liang, A. C., Feng, B., Zhao, J. J., Roberts, T. M., Mandel, G., Hannon, G. J., *et al.* (2005). A genetic screen for candidate tumor suppressors identifies REST. Cell *121*, 837-848.

## Investigation of Auto-Antibodies as Breast Cancer Biomarkers

To identify auto-antibodies that could be used as breast cancer biomarkers, we are generating a phage display library of coding fragments encompassing all open reading frames of the human genome. We designed approximately 467,900 overlapping oligonucleotides to cover 23,959 known protein-coding sequences. Each oligonucleotide encodes either 24 or 36 codons, with a 21-nucleotide region of overlap between each consecutive fragment. Twenty-two pools of 18,000-22,000 oligonucleotides (called HsORF pools for Homo *sapiens* open reading frame pools) were amplified by PCR and cloned into the Eco RI and Xho I sites of the lambda vector λSM2 (Silva, Li et al. 2005) to generate a λSM2-HsORF library before cloning the ORF fragments into vectors for phage display.

In order to display these coding fragments on phage, we chose two phage display vectors to allow for both lytic and filamentous versions of the library. For lytic phage display, we chose the T7-based vector T7Select®10-3b (Novagen). T7Select® 10-3b is a mid-copy vector that displays between 5-15 copies on the surface of the T7 capsid. The natural T7 10 capsid protein is produced in two forms (10A and 10B), where 10B is produced by a translational frameshift of 10A and comprises about 10% of the protein that comprises the capsid (Condron, Atkins et al. 1991). The T7Select® 10-3b vector makes only the 10B version of the capsid protein, and the coding sequences for the displayed peptides are cloned into a multiple cloning site to produce C-terminal fusions of the peptides of interest with 10B (Figure 1a). Phage from the T7Select® 10-3b vector are amplified on a bacterial host that carries an ampicillin-resistant plasmid expressing additional 10A capsid protein from a T7 promoter.

We modified T7Select® 10-3b by inserting a sequence encoding a FLAG epitope in the Not I and Xho I sites to allow for selection of in-frame peptides (Figure 1b). Once the T7FNS2 vector had been generated, the HsORF pools were were cloned into it to produce T7FNS2-HsORF pools (Figure 1d). This cloning results in a C-terminal fusion of the ORF fragments with the T7 10B capsid protein, followed by a C-terminal FLAG epitope. We have successfully cloned and validated eight of the twenty-two HsORF pools in the T7FNS2 vector. Furthermore, we have amplified phage from these eight T7FNS2-HsORF pools. Cloning and validation of the remaining 14 HsORF pools into the T7FNS2 vector is ongoing.

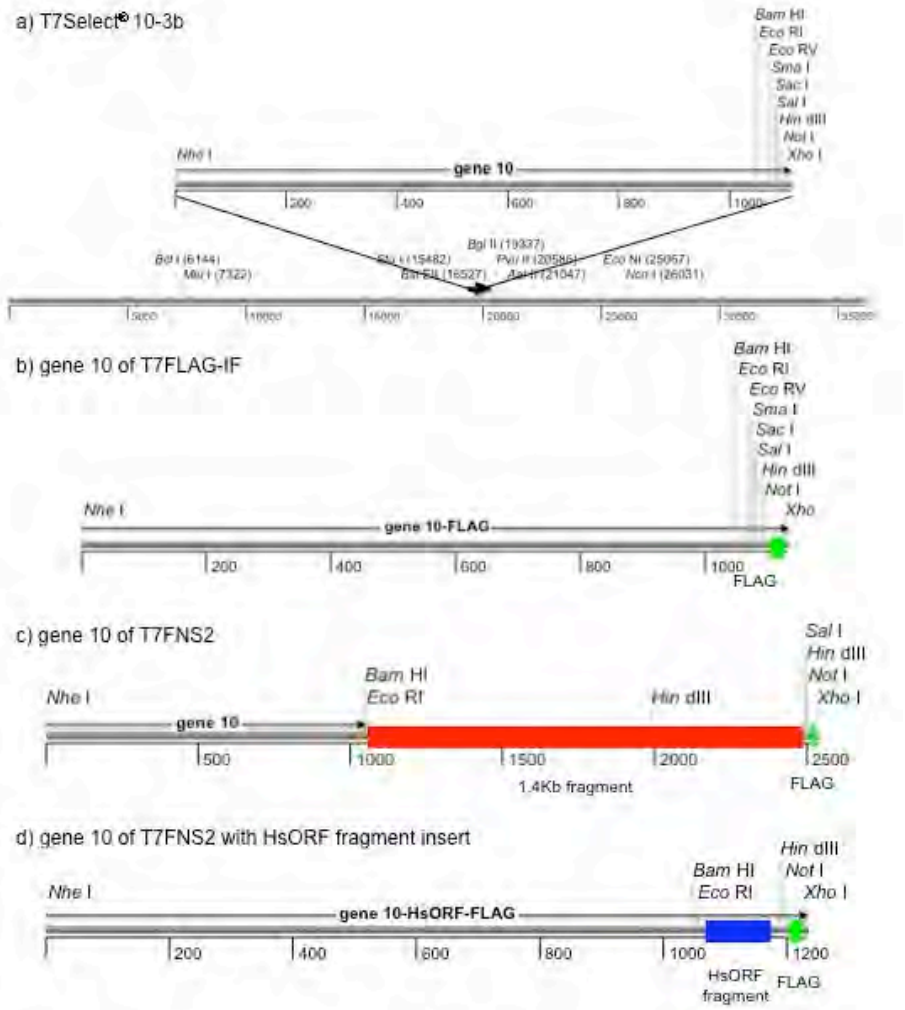



**Figure 1.**
T7FNS2 Phage Display Vector. a) The T7Select® 10-3b phage display vector from Novagen allows for C-terminal fusion of peptides to the 10B capsid protein. b) The T7Select® 10-3b vector was modified to generate T7FLAG-IF (as well as T7-FLAG+1n and +2n, not shown) by insertion of a FLAG epitope coding sequence in the Not I and Xho I sites. c) T7FNS2 was generated from T7FLAG+1n by insertion of a 1.4Kb fragment in the Bam HI and Sal I sites to accommodate in-frame HsORF fragments. d) HsORF fragment pools were cleaved from λSM2 by Eco RI and Xho I and inserted into the Eco RI and Sal I sites of T7FNS2 to form a 10B fusion followed by a FLAG epitope. To assess if we could use the C-terminal FLAG epitope to select for peptides expressed in-frame with the 10B capsid protein, we performed nitrocellulose plaque lifts and immunoblotting of T7Select® 10-3b parental phage and T7FLAG-IF phage (Figure 2). Immunoblotting of the membrane with a mouse anti-FLAG antibody conjugated with HRP revealed that the FLAG epitope was displayed on the surface of the T7FLAG-IF phage, but not of the T7Select® 10-3b parental phage. Furthermore, using a mouse anti-FLAG conjugated protein G sepharose resin and FLAG peptide, we were able to immunoprecipitate and elute at least 10-fold more FLAG-displaying phage than T7Select 10-3b phage, which do not display a FLAG epitope (data not shown). These data suggest that peptides can be displayed on the surface of T7 using this system and that the additional FLAG epitope can be used to select for in-frame peptides.

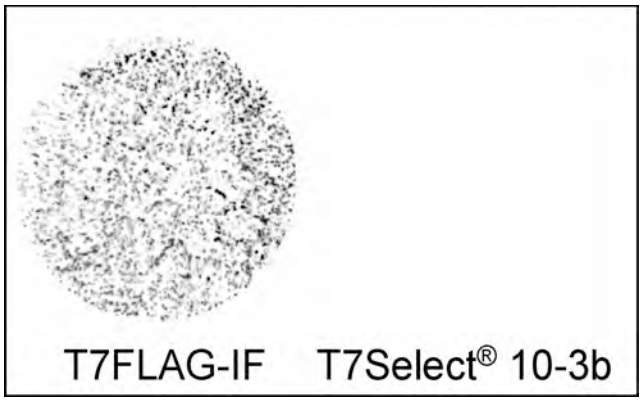



**Figure 2.** Phage from T7FLAG-IF Display a FLAG Epitope. T7FLAG-IF and T7Select® 10-3b phage were plated on BLT5403 host bacteria and incubated at 37°C until plaques developed. Plaques were lifted from each plate onto a nitrocellulose membrane. The membrane was blocked with 3% milk in TBST, immunoblotted with a mouse anti-FLAG antibody conjugated with HRP, and detected with ECL.

In addition, we modified the M13-based phagemid vector pPAO2 (Zacchi, Sblattero et al. 2003), to allow for filamentous phage display of the ORF fragment library. The pPAO2 phagemid vector allows for N-terminal fusion of peptides to pIII of M13. This vector is unique in that it allows for selection of in-frame peptides by conferring ampicillin resistance through B-lactamase expression. Protein coding sequences are cloned upstream of a B-lactamase and gene 3 fusion construct with the B-lactamase gene flanked by lox recombination sites. After cloning and selection of in-frame peptides by ampicillin resistance selection, the B-lactamase gene can be removed by Cre recombinase, leaving a standard M13 phage display vector with the peptide coding sequences fused to gene 3.

.

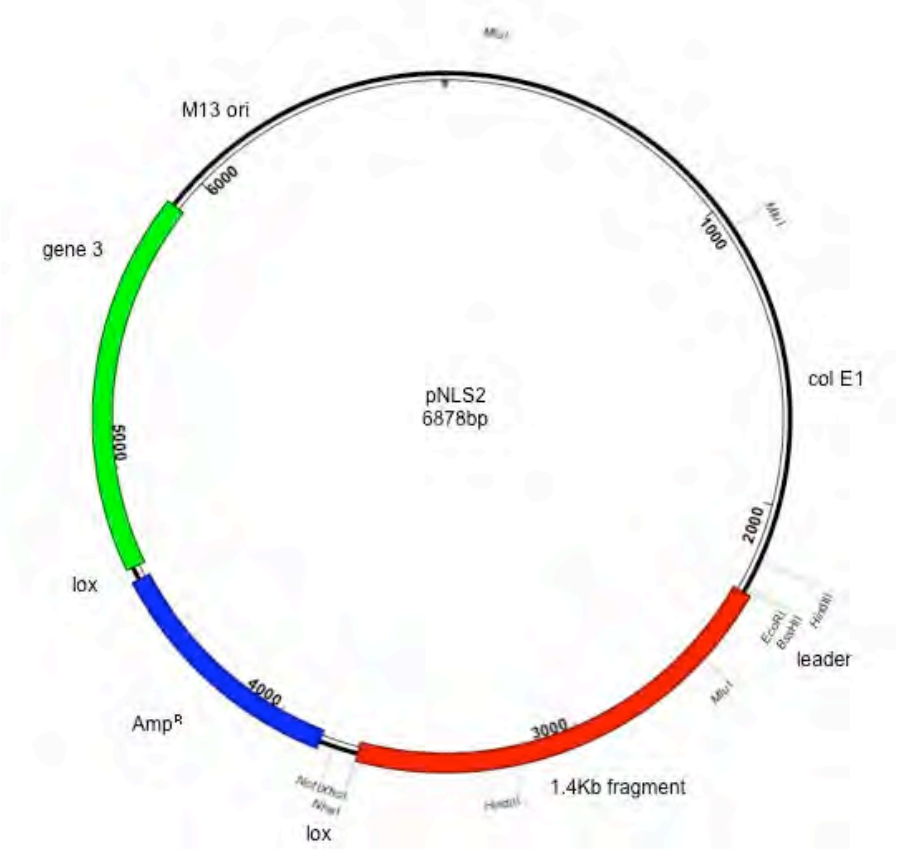



**Figure 3.** pNLS2 M13 Phagemid Vector. pNLS2 was generated by filling in the Eco RI site found after gene 3 in pPAO2 and by inserting a 1.4Kb fragment with Eco RI and Xho I sites into the Bam HI and Nhe I sites of pPAO2. HsORF fragment pools will be cloned into the Eco RI and Xho I sites to produce an N-terminal fusion with B-lactamase and pIII. The B-lactamase will be removed by Cre recombinase following selection of in-frame peptides to produce an N-terminal fusion with pIII.

The T7 and M13 libraries of phage-displayed peptides will be screened for binding to antibodies present in patient sera. This technology, coupled with microarray analysis, will be used to determine the peptide specificity for the entire antibody repertoire in a given serum sample. By comparing the immunoreactive peptide profiles of serum samples from multiple breast cancer patients and excluding peptides observed in normal samples, we aim to identify tumor-associated biomarkers. We are currently collaborating with the Breast SPORE blood bank at the Dana-Farber Cancer Institute to obtain serum samples for this project.

Condron, B. G., J. F. Atkins, et al. (1991). "Frameshifting in gene 10 of bacteriophage T7." J Bacteriol **173**(21): 6998-7003.

Silva, J. M., M. Z. Li, et al. (2005). "Second-generation shRNA libraries covering the mouse and human genomes." Nat Genet **37**(11): 1281-8.

Zacchi, P., D. Sblattero, et al. (2003). "Selecting open reading frames from DNA." Genome Res **13**(5): 980-90

## Key research accomplishments

1. Generation of a shRNA version 2 library in a high knockdown MSCV vector.
2. Screening of this library to identify more genes that repress transformation in mammary epithelial cells.
3. Identification of several genes including PTP14 phosphatase among 10 others that can suppress transformation of mammary cells. This was using our version 2 library and we are only part way through this screen.
4. Identification of several genes that when overexpressed cause transformation of mammary cells.
5. Demonstration that our new version 2 libraries are capable of being screened to identify hairpins that are toxic to cancer cells. This is a key finding essential to accomplishing the goals of this grant.
6. Generation of genome wide oligos that cover the entire coding capacity of the human genome.

With respect to the Statement of Work, we have accomplished many of the goals of year 1, 2 and 3 shown below.

## Statement of Work

**Year 1.**

**Task 1** (Months 1-12)

In the first year we anticipate beginning to work out the conditions for using the bar coding method to follow retroviruses containing hairpins as mixtures in complex libraries. We now have a library of 22,000 hairpins covering about 8,000 genes. We will be performing exploratory screens and optimizations to test the quality of the barcoding method. We must have this method working well to carryout the synthetic lethal screens.

We accomplished this goal in two ways. The first is we performed a bar code screen for potential tumor suppressors and identified several genes described in our first report and in Westbrook et al, 2005). Secondly, we have improved our vectors to allow single copy knockdown as described in Stegmeier et al. 2005. This is absolutely essential for the bar coding experiments we have proposed to kill cancer cells.

**Task 2** (Months 1-24 and possibly longer, an ongoing effort)

We will continue to expand the library during this period to encompass more genes. This will be done in collaboration with Dr. Greg Hannon.

We have accomplished this goal by the generation of a second generation library in the mir30 context as described in Silva et al., 2005. This covers 140,00 human and mouse shRNAs as was described in last years report. We have also developed new and better knockdown vectors to allow us to knock down genes with greater penetrance. Right now we feel we have nearly genome-wide coverage and are working on a new library which if successful will be a much better and more trustworthy library.

**Task 3** (Months 6-24)

We also will begin the process of analyzing the human genome for coding sequences to set up the bio-informatics analysis to generate a list of sequences we wish to express to look for auto-antibodies. We should begin synthesizing oligo nucleotides to cover human genes.

We have performed the analysis and designed oligonucleotides to cover the human genome. We have generated the oligonucleotides and are half way through cloning them in phage display vectors. Thus, we will NOT be purchasing a microarrayer as originally proposed. Instead we will use the funds for the microarrayer to purchase microarrays to deconvolute out immunoprecipitations.

**Year 2.**

**Task** 4 (Months 13-24)

In this period we plan to begin to carryout screen to look for genes which when knocked down by shRNA will interfere with the growth of cells containing defined mutations that lead to breast cancer. We will start with known tumor suppressors such as loss of p53 and Rb. We will use the barcoding methods. We may also screen for genes that sensitize cells to killing by gamma IR.

We have begun this aim and have screened one cell line for approximately half of our library. We have now started two other breast cancer cell lines in which to look for lethality. We have also initiated a synthetic lethal screen with cells with activated ras and begun to look for mediators of chemotheraputic drug resistance.

**Task 5** (Months 18-36)

We will begin to synthesize shRNA clones corresponding to the mouse genome.

We have gone a long way toward accomplishing this aim, We now have about 100,000 sequence verified shRNAs in our version 2 library already.

**Task 6** (Months 12-24)

We will expand the library of short coding regions for the autoantibody project and work out conditions to express these protein fragments in bacteria in a high through-put fashion.

We are ready to begin this task but have only just cloned the oligos which we had synthesized on microarrays.

**Year 3.**

**Task 7** (Months 24-36)

We will continue to screen for synthetic lethals with tumor causing mutations relevant to breast cancer. In addition, by this time we will be retesting the synthetic lethal positives from the initial screens performed in year two.

We are in the process of performing straight lethal and synthetic lethal experiments with ras. We have also tried to perform synthetic lethals with PTEN deficient cells but we did not find any synthetic lethals with PTEN. We have several good candidates for ras. We will continue to finish these screens.

**Task 8** (Months 24-36)

We plan to work out the conditions for placing the proteins expressing short segments of human proteins for the auto-antibody screening project onto glass slides for screening purposes.

We have abandoned this aim in that we switched our approach to a phage display library which does not require glass slide. We are in the process of making our first comprehensive library in a T7 display vector.

**Task 9** (Months 24-36)

We will continue to characterize the mouse shRNA library.

We are characterizing the mouse library. It is being transferred into our best knockdown vector.

## Reportable outcomes

Westbrook, T.F., Stegmeier, F. and Elledge SJ. (2005) Dissecting cancer pathways and vulnerabilities with RNAi. Cold Spring Harb Symp Quant Biol. 70:435-44.

Chang, K., Elledge, S.J. and Hannon, G.J. (2006) Lessons from Nature: microRNA-based shRNA libraries. Nat Methods. 3:707-14.

## Conclusions

### Progress on barcode screening for essential genes

It is clear from our current studies that we have overcome the main problem with performing bar code screens which is getting sufficiently good knockdown from single copy vectors and being able to reproducibly measure their abundance in complex pools by microarray hybridization. This has solved the main hurdle we had when we started this analysis. Clearly there are improvements we need to make. One such improvement is improving the percentage of hairpins we can detect with bar4 codes. We are currently trying to use two barcodes on the vectors to capture the remaining 20% of hairpins. We are also looking to reduce our noise by using inducible knockdown vectors, but this is still in the experimental stage.

### Identification of genes that confer acquired resistance to EGFR inhibitors.

We are clearly in the early stages of this project but have reproducible hits that confer resistance upon cells. We will be continuing the genome wide screen and validating the hits with independent hairpins to establish on target effects from off target effects. Once we have identified true genes that restrain EGFR inhibitor resistance we will begin a molecular analysis of their functions in cells and other cancer lines.

### Identification of genes whose knockdown confers synthetic lethality with the Ras oncogene,

This project is showing some promise in identifying synthetic lethal genes. As with the other screens we will need to finish the genome wide screen and then validate and pursue candidates both in the parent line and in other ras driven cancers. Hopefully we can test some of these hairpins in in vivo mouse models of ras induced carcinogenesis in the future.

### Screens for Tumor suppressors using the RNAi library

This ongoing effort should be completed in the next year. We ran into trouble with out new library as it was much more potent than our original library and also gave much more background, possibly because it was not a SIN vector. This means it takes much more time to determine a real hit. However, we are well on out way to identifying many more tumor suppressor candidates in this assay in HMECs and other cellular transformation models.

In addition we are performing the same screens with retroviral ORFeome libraries which are the equivalent to normalized full length cDNA libraries. This should identify novel breast cancer oncogenes.

**Identification and characterization of a novel REST ubiquitin ligase.**

As you recall from the first report, we identified an novel tumor suppressor REST using our transformation assay. We have found that REST is turned over rapidly by a ubiquitin-dependent pathway. In theory by accelerating this event one could transform cells as it reduced REST function> Therefore we have searched for the ubiquitin ligase for REST using a genetic screen. We found that betaTRCP, and SCF F-box protein is responsible for REST degradation. We went on to show that overproduction of this F-box protein can transform HMECs. We are now testing whether this is REST-dependent.

**Investigation of Auto-Antibodies as Breast Cancer Biomarkers**

We are still in early stages of this project but we are well on our way to having the libraries in hand and the patient samples and should have some data on the method by the next progress report.

**References**


Chong, J. A., Tapia-Ramirez, J., Kim, S., Toledo-Aral, J. J., Zheng, Y., Boutros, M. C., Altshuller, Y. M., Frohman, M. A., Kraner, S. D., and Mandel, G. (1995). REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. Cell *80*, 949-957.

Condron, B. G., J. F. Atkins, et al. (1991). "Frameshifting in gene 10 of bacteriophage T7." J Bacteriol **173**(21): 6998-7003.

Debnath, J., and Brugge, J. S. (2005). Modelling glandular epithelial cancers in three-dimensional cultures. Nat Rev Cancer *5*, 675-688.

Debnath, J., Mills, K. R., Collins, N. L., Reginato, M. J., Muthuswamy, S. K., and Brugge, J. S. (2002). The role of apoptosis in creating and maintaining luminal space within normal and oncogene-expressing mammary acini. Cell *111*, 29-40.

Hynes, N. E., and Lane, H. A. (2005). ERBB receptors and cancer: the complexity of targeted inhibitors. Nat Rev Cancer *5*, 341-354.

Mukohara, T., Engelman, J. A., Hanna, N. H., Yeap, B. Y., Kobayashi, S., Lindeman, N., Halmos, B., Pearlberg, J., Tsuchihashi, Z., Cantley, L. C., Tenen, D. G., Johnson, B. E., and Janne, P. A. (2005). Differential effects of gefitinib and cetuximab on non-small-cell lung cancers bearing epidermal growth factor receptor mutations. J Natl Cancer Inst *97*, 1185-1194.

Muthuswamy, S. K., Li, D., Lelievre, S., Bissell, M. J., and Brugge, J. S. (2001). ErbB2, but not ErbB1, reinitiates proliferation and induces luminal repopulation in epithelial acini. Nat Cell Biol *3*, 785-792.

Schoenherr, C. J., and Anderson, D. J. (1995). The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. Science *267*, 1360-1363.

Shirasawa, S., Furuse, M., Yokoyama, N., and Sasazuki, T. (1993). Altered growth of human colon cancer cell lines disrupted at activated Ki-ras. Science *260*, 85-88.

Silva, J. M., Li, M. Z., Chang, K., Ge, W., Golding, M. C., Rickles, R. J., Siolas, D., Hu, G., Paddison, P. J., Schlabach, M. R., *et al.* (2005). Second-generation shRNA libraries covering the mouse and human genomes. Nat Genet *37*, 1281-1288.

Westbrook, T. F., Martin, E. S., Schlabach, M. R., Leng, Y., Liang, A. C., Feng, B., Zhao, J. J., Roberts, T. M., Mandel, G., Hannon, G. J., *et al.* (2005). A genetic screen for candidate tumor suppressors identifies REST. Cell *121*, 837-848.

Zacchi, P., D. Sblattero, et al. (2003). "Selecting open reading frames from DNA." Genome Res **13**(5): 980-90

**Appendices**

None